DOCUMENT RESUME

ED 176 707 HE 011 884

AUTHOR Aleamoni, Lawrence M.
TITLE Are There Differences in Perceived Teaching Effectiveness Between Males and Females in Anthropology?
PUB DATE 16 Nov 78
NOTE 13p.; Paper presented at the Annual Meeting of the American Anthropological Association (Los Angeles, California, November 16, 1978); not available in paper copy due to marginal legibility of original

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Anthropology; *College Faculty; College Freshmen; Effective Teaching; Females; Higher Education; Males; *Sex Differences; *Student Evaluation of Teacher Performance; *Student Opinion; Universities; *Women Teachers

ABSTRACT

The differences in student ratings between male and female instructors in Anthropology were examined in a study of approximately 3,156 student ratings of 82 male and 32 female instructors over a two-year period. An assessment instrument was developed to assess teaching effectiveness. Analysis of the results seemed to indicate that instructors in Anthropology were rated differently by their students in spite of extensive research findings to the contrary. On careful examination of the data it was discovered that the female instructors were teaching a disproportionately higher number of courses at the freshman level than their male counterparts. Since there had been extensive research evidence reporting more favorable ratings being assigned by graduate and/or upper division students the data in the present study needed further examination. When the ratings for the male and female instructors were reexamined within each course level all significant differences disappeared except at the freshman level, where the males were rated significantly more favorably than the females. It is noted that there were more females teaching multiple sections of a single course at the freshman level than males, which may account for the rating variance. (Author/PHR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED176707

ARE THERE DIFFERENCES IN PERCEIVED TEACHING EFFECTIVENESS BETWEEN MALES AND FEMALES IN ANTHROPOLOGY?

Lawrence M. Aleamoni

University of Arizona

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Lawrence M. Aleamoni

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

HE 011 884

Abstract

A review of the research literature on differences in student ratings of male and female instructors resulted in the conclusion that to date there have been no significant differences reported. The present study compared student ratings between male and female instructors in Anthropology at the University of Arizona and found significant differences in ratings at the freshman level. Implications of the study are discussed.

Perceived Teaching Effectiveness in Anthropology

Footnote

The results of this study were presented at the annual meeting of the American Anthropological Association in Los Angeles on November 15, 1978.

ARE THERE DIFFERENCES IN PERCEIVED TEACHING EFFECTIVENESS BETWEEN MALES AND FEMALES IN ANTHROPOLOGY?

In recent years the use of student evaluations of course and instructors has increased rapidly on college campuses. These evaluations are used typically to provide feedback to instructors on their strengths and weaknesses in the classroom and sometimes in faculty promotion and tenure decisions. This use of student evaluations of instruction has generated a great deal of interest in student, instructor, and course characteristics which may affect these evaluations.

A considerable body of research has accumulated relating student and course characteristics to student ratings of instructors. For a detailed review of this research see Costin, Greenough, and Menges (1971) and Kulik and Kulik (1974). Due in part to the use of different course evaluation forms and to the use of differing research methodologies, the results of these investigations are often discrepant.

Conflicting results, for example, have been obtained when relating student sex to students evaluations of instruction: Goodhartz (1948), Isaacson, McKeachie, Milholland, Lin, Hofeller, Baerwaldt, and Zinn (1964), Hildebrand, Wilson and Dienst (1971), Doyle (1974), and Centra and Creech (1976) reported no differences between faculty ratings made by male and female students. In addition, several studies reported no differences in overall ratings of instructors made by male and female students, or in the ratings received by male and female instructors (Bendig 1953; Caffrey 1969; Downie 1952; Harris 1975; Heilman and Armentrout 1936; Lovell and Haner 1955; Remmers 1939). Conversely, Bendig (1952) found female students to be more critical of male instructors

than their male counterparts; and more recently Walker (1969) found that female students rated female instructors significantly higher than they rated male instructors. In addition, several investigators (Ashton 1975; Elmore and La Pointe 1974; Kohlan 1973; McKeachie, Lin, and Mann 1971; Pohlmann 1975) reported that female students rate instructors higher on some subscales of instructor evaluation forms than do male students.

The current study was designed to examine the differences in student ratings between male and female instructors in Anthropology at a large southwestern university. Based upon the previously cited research it is hypothesized that there will be no significant differences between the ratings of male and female instructors in Anthropology.

Method

The Arizona Course/Instructor Evaluation Questionnaire (CIEQ), an instrument used to assess teaching effectiveness via student ratings, was developed through the use of logical grouping of items and factor analysis. The instrument is basically divided into four sections: (a) a student, course, and coding information section, (b) a standard item section, (c) an optional item section, and (d) an open-ended response section. This study will deal with only the standard item section.

In the student information section students are asked to respond to questions concerning their status, whether they are taking the course for pass/fail versus standard grading, whether they are taking the course as a required or an elective, their sex, their expected grade, and whether the course is within their major or minor. In the course information section the students are asked to rate three

single general items regarding the course content, major instructor, and course in general on a six-point scale.

The standard item section contains 21 items which form five subscales (general course attitude, method of instruction, course content, interest-attention, and instructor) and a total. To each item students indicate their degree of agreement or disagreement on a four-point scale of agree strongly (AS), agree (A), disagree (D), and disagree strongly (DS). Since items are phrased both positively and negatively, a weight of four is assigned to the most favorable response down to a weight of one for the least favorable response. See Aleamoni (1977) for a detailed discussion of the reliability and validity studies conducted with the CIEQ.

Subjects

The units of study were 114 graduate and undergraduate courses at the University of Arizona. These units represented approximately 3,156 students rating their instructors and courses using the CIEQ. Course evaluation questionnaires were completed by students for 82 male and 32 female instructors from Spring, 1976 through Spring, 1978.

Procedure

Analyses of variance were run between the male and female instructors using each of the five CIEQ subscales and the total as dependent measures. In order to determine if course level might be associated with the instructor's sex a chi-square analysis was run. In addition, where significant F-values were obtained t-tests were then run between males and females stratified by course level.

Results

The analysis of variance results along with the means and standard deviations for the two groups are presented in Table 1. There were significant differences between the male and female ratings on each of the CIEQ subscales but not the total, with the males receiving the more positive ratings.

TABLE 1

Means, Standard Deviations, F-ratios, and Probabilities for the 82 Male and 32 Female Instructors on the Six CIEQ Subscales and Total

Dependent Variable	Males		Females			
	Mean	S.D.	Mean	S.D.	F-ratio	Probability
General Course Attitude	3.1880	.3540	2.9070	.3582	14.4081	.0002
Method of Instruction	2.8830	.3721	2.6590	.3613	8.4755	.0043
Course Content	2.9611	.2815	2.7881	.2787	8.7414	.0038
Interest-Attention	2.8833	.3269	2.6794	.3519	8.5779	.0041
Instructor	3.2348	.2893	3.1034	.2861	4.7776	.0309
Total	3.0348	.6300	2.8366	.2916	2.9113	.0907

However, in an earlier study by Aleamoni and Graham (1974) it was determined that if such ratings were stratified by course level (freshman, sophomore, junior, senior, and graduate) no such differences would exist. In order to determine if there was a disproportionate number of males versus females teaching courses at the various levels a chi-square analysis was run yielding a chi-

square of 14.91 which with 4 degrees of freedom indicates that there is a definite lack of proportional representation of males versus females teaching at each course level.

Independent t-tests were then run between males and females teaching at each course level for the CIEQ Instructor subscale yielding the data reported in Table 2. These results indicate that males and females are rated differently in only the freshman level courses.

TABLE 2

Means, Standard Deviations, t-tests, and Probabilities for the Male and Female Instructors Stratified by Course Level Under the CIEQ Instructor Subscale

Course Level	Males			Females			T-test	Probability
	N	Mean	S.D.	N	Mean	S.D.		
Freshman	18	3.2189	.2716	18	3.0200	.2710	2.1992	.0175
Sophomore	17	3.1382	.3657	4	3.3675	.2626	-1.1739	.125
Junior	14	3.1536	.3362	6	3.1117	.2883	.2654	.395
Senior	18	3.3567	.1392	2	3.2250	.3748	1.0934	.145
Graduate	15	3.1780	.2849	2	3.1800	.2828	-.0093	.497

Discussion

The initial analysis of variance results seemed to indicate that instructors in Anthropology are rated differently by their students when subscale scores are used in spite of the extensive research findings to the contrary. However, when the data was examined in more detail it was discovered that the female instructors were teaching a disproportionately higher number of courses at the freshman

level than their male counterparts. Since there has been extensive research evidence reporting significantly more favorable ratings being assigned by graduate and/or upper division students (Costin, Greenough, and Menges 1971; Kohlan 1973; Aleamoni and Graham 1974; Pohlmann 1975) the data in the present study needed further examination.

When the ratings for the male and female instructors were re-examined within each course level all significant differences disappeared except at the freshman level where the males were rated more favorably than the females. It is interesting to note that there were more females teaching multiple sections of a single course at the freshman level than males.

In summary, it appears that males and females teaching in Anthropology at a large southwestern university are not rated differently in terms of their instructional effectiveness except by students at the freshman level. Since the latter result does not seem to be supported by the preponderance of published research evidence, that difference may be a result of the heavier multiple section load of the females or the fact that they may have less of a choice of freshman level courses that they would like to teach.

Further studies need to be conducted in other departments of Anthropology to see if these results can be replicated.

References

Aleamoni, L. M. The Arizona course/instructor evaluation questionnaire: A description of its development (4R3). Tucson: University of Arizona, Office of Instructional Research and Development, 1977.

Aleamoni, L. M., & Graham, M. H. The relationship between CEQ ratings and instructor's rank, class size, and course level. Journal of Educational Measurement, 1974, 11, 189-202.

Ashton, R. H. Correlations of rates of teaching effectiveness: Gordon's survey of interpersonal values. Psychological Reports, 1975, 36, 883.

Bendig, A. W. A preliminary study of the effect of academic level, sex, and course variables on student rating of psychology instructors. Journal of Psychology, 1952, 34, 2-126.

Bendig, A. W. Student achievement in introductory psychology and student ratings of the competence and empathy of their instructors. Journal of Psychology, 1953, 36, 427-433.

Caffrey, B. Lack of bias in student evaluations of teachers. Proceedings of the 77th Annual Convention of the American Psychological Association, 1969, 4, 641-642.

Centra, J. A., & Creech, F. R. The relationship between students, teacher, and course characteristics and student ratings of teacher effectiveness. Princeton, N. J.: Educational Testing Service, 1976.

Costin, F., Greenough, W. T., & Menges, R. J. Student ratings of college teaching: Reliability, validity, and usefulness. Review of Educational Research, 1971, 41 (5), 511-535.

Downie, N. W. Student evaluation of faculty. Journal of Higher Education, 1952, 23, 495-496, 503.

Elmore, P. B., & La Voie, K. A. Effects of teacher sex and student sex on the evaluation of college instructors. Journal of Educational Psychology, 1974, 66 (3), 386-389.

Goodhartz, A. W. Student attitudes and opinions relating to teaching at Brooklyn College. School and Society, 1948, 68, 340-349.

Harris, M. B. Sex role stereotype and teacher evaluations. Journal of Educational Psychology, 1975, 67, 751-756.

Heilman, J. D., & Armentrout, W. D. The rating of college teachers on ten traits by their students. Journal of Educational Psychology, 1936, 27, 197-216.

Hildebrand, M. Wilson, R. C., & Dienst, E. R. Evaluating university teaching. Berkeley, California: University of California, Center for Research and Development in Higher Education, 1971.

Isaacson, R. L., McKeachie, W. J., Milholland, J. E., Lin, Y. G. Hofeller, M., Baerwaldt, J. W., & Zinn, K. L. Dimensions of student evaluations of teaching. Journal of Educational Psychology, 1964, 55, 344-351.

Kohlan, R. G. A comparison of faculty evaluations early and late in the course. Journal of Higher Education, 1973, 44, 587-595.

Kulik, J. A., & Kulik, C. C. Student ratings of instruction. Teaching of Psychology, 1974, 1 (2), 51-57.

Lovell, G. D., & Haner, C. F. Forced-choice applied to college faculty rating. Educational and Psychological Measurement, 1955, 15, 291-304.

McKeachie, W. J., Lin, Y., & Mann, W. Student's ratings of teacher effectiveness: Validity studies. American Educational Research Journal, 1971, 8, 435-445.

Pohlmann, J. T. A multivariate analysis of selected class characteristics and student ratings of instruction. Multivariate Behavioral Research, 1975, 10, 81-91.

Remmers, H. H. Appraisal of college teaching through ratings and student opinion. In 27th Yearbook of the National Society of College Teachers of Education. Chicago: University of Chicago Press, 1939.

Walker, B. D. An investigation of selected variables relative to the manner in which a population of junior college students evaluate their teachers. Dissertation Abstracts, 1969, 29 (9-B), 3474.